Pergamon

Computers Math. Applic. Vol. 34, No. 1, pp. 97–102, 1997

0898-1221/97 $17.00 + 0.00

PII: S0898-1221(97)00101-6

An ODE Method of Solving Nonlinear Programming

ZONGFANG ZHOU
Department of Basic Studies, Chongqing University of Posts and Telecommunication
630065, P.R. China

Y. SHI
Department of Information Systems and Quantitative Analysis
University of Nebraska at Omaha, Omaha, NE 68182, U.S.A.

(Received July 1996; accepted August 1996)

Abstract—This paper proposes a new method to solve general constrained optimization problem. The problem of finding the local optimal points of a nonlinear programming problem with equality and inequality constraints is considered by solving the ODE (i.e., ordinary differential equation) with an appropriate numerical procedure. Moreover, the rate of convergence to optimal points is quadratic. Some numerical result is given to show the efficiency of the proposed method.

Keywords—Constrained optimization, ODE, Local optimal points.

1. INTRODUCTION

In this paper, a new method is presented for dealing with general constrained optimization problem by solving ordinary differential equation with an appropriate numerical procedure. In this section, the constrained optimization problem is changed into solving singular solutions of an autonomous system. In Section 2, a decline line differential equation is introduced. It is proved that not only the right trajectory of the solutions of the system about part variables is a decline line of object function at initial points, but also the limit point of the trajectory is a Kuhn-Tucker point of the original problem. Section 3 presents a numerical procedure for the method. The concluding remark is given in Section 4.

The general constrained optimization problem is considered

$$\begin{cases} \min f(x), \\ x \in \bar{S} = \{x \in R^n \mid h(x) = 0,\ g(x) \leq 0\}, \end{cases} \tag{1}$$

where

$$f(x) \in C^2, \qquad h(x) = (h_1(x), \dots, h_m(x))^\top, \qquad g(x) = (g_1(x), \dots, g_r(x))^\top.$$

After introducing a new slack variable Z, we obtain the equality constrained optimization problem [1]:

$$\begin{aligned} &\min F(y) = f(x), \\ &y \in S = \left\{ y \in E^{n+r} \mid H(y) = \begin{pmatrix} h(x) \\ g(x) + \dfrac{Z^2}{2} \end{pmatrix} = 0 \right\}. \end{aligned} \tag{2}$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Let $I = \{i \mid g_i(x) = 0,\ i = 1,2,3,\dots,r\}$ be the take effect constrained index system of (1). Then, we have

$$g_I = \{g_i, i \in I\}, \qquad Z_I = \{Z_i, i \in I\},$$
$$\bar{H}(y) = \begin{bmatrix} h(x) \\ g_I(x) + \dfrac{Z_T^2}{2} \end{bmatrix},$$
$$A = \nabla H(y) = \begin{pmatrix} \nabla h^\top(x) & 0 \\ \nabla g^\top(x) & ZEr \end{pmatrix},$$
$$\bar{A} = \nabla \bar{H}(x) = \begin{pmatrix} \nabla h^\top(x) & 0 & 0 \\ \nabla g^\top(x) & Z_I E & 0 \end{pmatrix},$$

where $ZE_r = \operatorname{diag}(Z_1,\dots,Z_r)$, $Z_I E = \operatorname{diag}(z_i)$.

We assume that the vectors $(\nabla h_i(x), \nabla g_j(x),\ i = 1,2,\dots,m;\ j = 1,2,\dots,r)$ are linearly independent. To solve the above problem, we introduce the following system of ordinary differential equation (ODE):

$$\begin{aligned} \begin{bmatrix} B_n \dot{X} \\ B_r \dot{X} \end{bmatrix} + \begin{pmatrix} \nabla h(X)\lambda m + \nabla g(X)\lambda r \\ ZEr\lambda r \end{pmatrix} &= \begin{pmatrix} -\nabla f(X) \\ 0 \end{pmatrix}, \\ \begin{bmatrix} \nabla h^\top(X)\dot{X} \\ \nabla g^\top(X)\dot{X} + ZEr\dot{Z} \end{bmatrix} &= \begin{bmatrix} -h(X) \\ -g(X) - \dfrac{Z^2}{2} \end{bmatrix}. \end{aligned} \tag{3}$$

Statement 1. (See [2].) If

$$B = \begin{bmatrix} B_n & 0 \\ 0 & B_r \end{bmatrix},$$

is of full rank, then $AB^{-1}A^\top$ and

$$D = \begin{pmatrix} B & A^\top \\ A & 0 \end{pmatrix},$$

are nonsingular, where $A = \nabla H(y)$.

The proof of the above statement can be easily shown. We can solve (3) for y and λ uniquely and obtain

$$\dot{y} = \Phi(y) = -PB^{-1}\nabla F - \bar{P}H, \tag{4}$$
$$\lambda = -\left(AB^{-1}A^\top\right)^{-1} AB^{-1}\nabla F + \left(AB^{-1}A^\top\right)^{-1} H, \qquad \text{where} \tag{5}$$
$$P = E - B^{-1}A^\top \left(AB^{-1}A^\top\right)^{-1} A,$$
$$\bar{P} = B^{-1}A^\top \left(AB^{-1}A^\top\right)^{-1}.$$

Definition 1. A point $\bar{X} \in \bar{S}$ is called a critical point of (1), if there exist (unique) real numbers $\bar{\lambda}_i$, $\bar{\mu}_j$, $i \in I(\bar{x})$, $j = 1,2,\dots,m$ satisfying

$$\begin{aligned} \nabla f(\bar{x}) + \sum \bar{\mu}_j \nabla h_j(\bar{x}) + \sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x}) &= 0, \\ \bar{\lambda}_t \cdot g(\bar{x}) &= 0. \end{aligned} \tag{6}$$

The numbers $\bar{\lambda}_i, \bar{\mu}_j$ are called Lagrange parameters. A critical point $\bar{x}$ is a Kuhn-Tucker point if $\bar{\lambda}_i \geq 0$ for all $i \in I(\bar{x})$.

Theorem 1.

(i) *If $y^* = (x^*, z^*)^\top$ satisfies $\Phi(y^*) = 0$, then y^* is a critical point of problem (2).*
(ii) *If y^* is a critical point of problem (2), then x^* is a critical point of problem (1).*
(iii) *If x^* is a critical point of problem (1), then $y^* = (x^*, \sqrt{-2g(x^*)}\,)^\top$ satisfies $\Phi(y^*) = 0$.*

The above theorem shows that we can obtain a Kuhn-Tucker point if $\lambda_r \geq 0$ in (3). In the sequel we discuss the second-order sufficient condition for local optimality. Let

$$M(y) = \{u \mid Au = 0, u = (v, w) \in R^{n+r}\}$$

be the tangent plane of feasible sets at y, and

$$\bar{M}(y) = \{u \mid Au = 0, u = (v, w_1, w_2) \in R^{n+r}\},$$

where $v \in R^n$, $w_1 \in R^{[I]}$, $w_2 \in R^{r-[I]}$, and $[I]$ is the number of elements in the index set I. Then,

$$M(x) = \{u \mid h^\top(x)u = 0, g_I^\top(x)u = 0, u \in R^n\}$$

is the tangent plane of feasible set $\bar{S}$ at x, and

$$\bar{M}^0(y) = \{u \mid Au = 0, u = (v, 0)^\top \in R^{n+1}\}.$$

This leads to the next corollary.

COROLLARY 1.

$$\begin{array}{ll} (i) & u = (v,w)^\top \in \bar{M}(y) \Rightarrow v \in M(x), \\ (ii) & u = (v,w)^\top \in M(y) \Rightarrow v \in M(x), \\ (iii) & u = (v,w)^\top \in \bar{M}^0(y) \Rightarrow v \in M(x). \end{array}$$

Now, let $L(y) = H^\top(y)\lambda + F(y)$ and $l(x) = f(x) + h^\top(x)\lambda_m + g^\top(x)\lambda_r$. Then, we see that

$$\nabla^2 L(y) = \begin{bmatrix} \nabla^2 l(x) & 0 \\ 0 & \sum_{i=1}^{T} \lambda_T^{(i)} Er \end{bmatrix}.$$

This provides the following theorem.

THEOREM 2. *Let $y^* = (x^*, z^*)^\top$ satisfy $\Phi(y^*) = 0$, and $\lambda_r \geq 0$ in (5), If $\nabla^2 L(y)$ is positive (negative) definite on the tangent plane $\bar{M}(y^*)$, then x^* is a strict local minimal (maximal) point of (1).*

2. DECLINE LINE DIFFERENTIAL EQUATIONS

We denote the solution of the system (3) with initial condition $y = \xi$, at $t = 0$ by $y(0) = \xi$, and the whole trajectory $C(\xi) = \{y(t, \xi; 0), t \in T\}$. Then $T(\xi) = (a(\xi), b(\xi))$ is denoted as a maximal interval of existence of the solution.

DEFINITION 2. Let K be a subset of R^n, if $C(\xi) \subset K$ for any $\xi \in K$, then the set K is called an invariant set.

According to Definition 2, we can obtain the following results [3].

COROLLARY 2. *The set S of (2) is invariant.*

COROLLARY 3. *If $y = (x, z)^\top \in S$, then $x \in \bar{S}$, and $y = (x, \sqrt{-2g(x)}\,)^\top \in S$.*

COROLLARY 4. *$\bar{S}$ is invariant.*

From the above corollaries, we can easily prove the following deduction.

DEDUCTION 1. If $y(t, \xi; 0)$ and $x(t, \xi; 0)$ have positive limit set $\Gamma^+(\xi)$ and $\Gamma^+(\xi_x)$, respectively, then

$$\Gamma^+(\xi) \subset S \quad \text{and} \quad \Gamma^+(\xi_x) \subset \bar{S}.$$

According to the results of [4], we have the following definition.

DEFINITION 3. Let curve $y(t)$ be defined on the interval (α, β), $\alpha \le 0 \le \beta$, where $\alpha^2 + \beta^2 \ne 0$, and $y(t)$ belong to definition region of $F(y)$; if $F(y(t))$ is strictly decreasing with respect to the variables $x(t)$ on (α, β), then $y(t)$ is called for the decline line of $F(y)$ with respect to variables $x(t)$ at $x(0) = x^0$. If

$$\lim_{t \to \beta} x(t),$$

exists and equals the minimal point of $F(y)$ with respect to the x variables, then $x(t)$ is called the regular decline line of $F(y)$ at $x(0) = x^0$.

Now, we consider the following initial problem:

$$\begin{aligned} \dot{y} = \Phi(y) &= -PB^{-1}\nabla F - \bar{P}H, \\ y(0) = y^0 &= \left(x^0, t^0\right)^\top. \end{aligned} \tag{7}$$

Denote $T(y^0) = (\alpha(y^0), \beta(y^0))$ the maximal interval where the solution of problem (7) exist, with $\alpha(y^0) < 0$, $\beta(y^0) > 0$.

THEOREM 3. *If $B(y)$ is positive definite on the $\bar{M}(y)$ for any $y \in S$, and $\nabla F(y^0) \ne 0$, then the solution $y(t)$ of (7) is the decline line of $f(x)$ at $x(0) = x^0$.*

PROOF. Since S is invariant, from the the assumption of the theorem, we have

$$\frac{dF(y)}{dt} = -\dot{y}^\top B \dot{y}.$$

Since $\dot{y} = \Phi(y) \ne 0$, we see

$$\nabla F(y) \ne 0,$$

and

$$\frac{dF(y)}{dt} = \frac{df(x)}{dt} < 0, \qquad t \in T\left(y^0\right). \tag{8}$$

For any $t \in T(y^0)$, we have $\nabla F(y) \ne 0$. If there exist $\bar{t} \in T(y^0)$ and $\bar{y} = y(\bar{t})$ satisfying $\nabla F(\bar{y}) = 0$, then the solution $y(t)$ of the initial problem

$$\frac{dy}{dt} = \Phi(y), \qquad y\left(\bar{t}\right) = \bar{y}, \tag{9}$$

exists.

But $y = \bar{y}$ is another constant solution of the problem. Therefore, neither $\nabla F(y^0) \ne 0$, nor $\nabla F(\bar{y}) = 0$ occur, and we can deduct that

$$y^0 \ne \bar{y}.$$

This contradicts the uniqueness of the solution. It follows that (8) is true, i.e., $F(y(t))$ along (9) is strictly decreasing on $T(y^0)$. ∎

The following theorem will guarantee partly the stability of our algorithms presented in Section 3 in a neighborhood of the solution.

THEOREM 4. *For the initial point $y^0 \in S$, let $B(y)$ be positive definite. If the set*

$$N(f, x^0) = \left\{x \mid x \in \bar{S}, f(x) \le f\left(x^0\right)\right\},$$

is a bounded closed set, and $\nabla f(x^0) \ne 0$, then the right trajectory of problem (7) is the decline line at $x(0) = x^0$. Furthermore, its limit point is a critical point of (1).

DEDUCTION 2. Under the assumption of Theorem 4, if $\bar{x}$ is a limit point of the right trajectory of problem (7), and $\lambda_r \ge 0$ in (5) at $x = \bar{x}$, then $\bar{x}$ is a Kuhn-Tucker point of problem (1).

3. A NUMERICAL PROCEDURE

Based on the above discussion, we propose a numerical procedure to solve a constrained nonlinear programming by ODE as follows.

Procedure 1

STEP 1. For any given problem, determine an initial point y^0 and $B(y)$. The form of $B(y)$ can be [5]:

$$B(y) = \nabla^2 L(y) = \nabla^2 F(y) + \lambda_0^{\top} \nabla^2 H(y), \tag{10}$$

where $\lambda_0 = (-AA^{\top})^{-1} A \nabla F^{\top}(y)$ is the first-order estimate of the Lagrange multiplier λ.

STEP 2. Once initial point y^0 and $B(y)$ are determined, find local optimal point of problem (1) by solving the critical point of the following initial problem

$$\begin{aligned} \frac{dy}{dt} &= \Phi(y) = -PB^{-1}\nabla F^{\top} - \bar{P}H, \\ y(T_0) &= y_0. \end{aligned} \tag{11}$$

There are many numerical integration methods to solve (11). We use the fourth-order Runge-Kutta method to integrate the trajectory of the following example for finding numerically its initial point. Then, the Euler method is used to accelerate convergence to the critical point. The convergence rate of this method is second-order. For suitable initial points, the convergence rate of this algorithm is more rapid than that of general optimization method.

The trajectory is supposed to converge to a critical point if the conditions

$$\begin{aligned} \|\Phi(y)\| &< \varepsilon, \\ \|H(y)\| &< \varepsilon, \end{aligned}$$

are satisfied.

Now, we show the following example where $\varepsilon = 0.001$, to illustrate the procedure.

EXAMPLE. We consider the following constrained nonlinear programming problem:

$$\begin{aligned} \min f(x_1, x_2) &= (x_1 - 2)^2 + (x_2 - 1)^2, \\ h(x_1, x_2) &= x_1 - 2x_2 = 0, \\ g(x_1, x_2) &= x_1 - 1 \le 0, \qquad x_1, x_2 \ge 0. \end{aligned}$$

Using the above procedure, we solve the problem and obtain three-level extreme points as the solutions in Table 1.

Table 1.

	$x^{(0)}$	$x^{(1)}$	$x^{(2)}$
$x^{(i)}$	1.0000000	1.0332840	0.9971844
$x_2^{(i)}$	0.4500000	0.4835770	0.5014097
$f\left(x_1^{(i)}, x_2^{(i)}\right)$	1.1125000	1.2014590	1.2542332

4. CONCLUDING REMARKS

We described a new numerical method for solving general constrained optimization problems. According to this method, we first transform a given problem into an ordinary differential equation (ODE) problem. Then, a numerical procedure is applied to solve the ODE problem. Finally, the Kuhn-Tucker point of the original problem is identified as the limit point of the trajectory of an ODE problem. This method has great possibilities in solving large-scale optimization problems in engineering or business decision making.

REFERENCES

1. K. Tanabe, A geometric method in nonlinear programming, *Journal of Optimization Theory and Applications* **2** (30), 181–185 (1980).
2. Zongfang Zhou, Dealing with the constraints of optimization by ODE method, *Journal of China Univ. of Posts and Telecom.* **1** (1), 65–67 (1994).
3. Zongfang Zhou, The applications of differential equations in optimization constraints, *Journal of Systems Engineering* **11** (6), 57–60 (1993).
4. Zongfang Zhou, Applications of the stability theory on the constrained optimization, *Journal of Chongqing Univ. of Posts and Telecom.* **3** (2), 71–78 (1991).
5. Zongfang Zhou, The discussion about the coefficient matrix in ODE method, *Journal of Systems Engineering and Electronics* **17** (9), 77–80 (1995).